

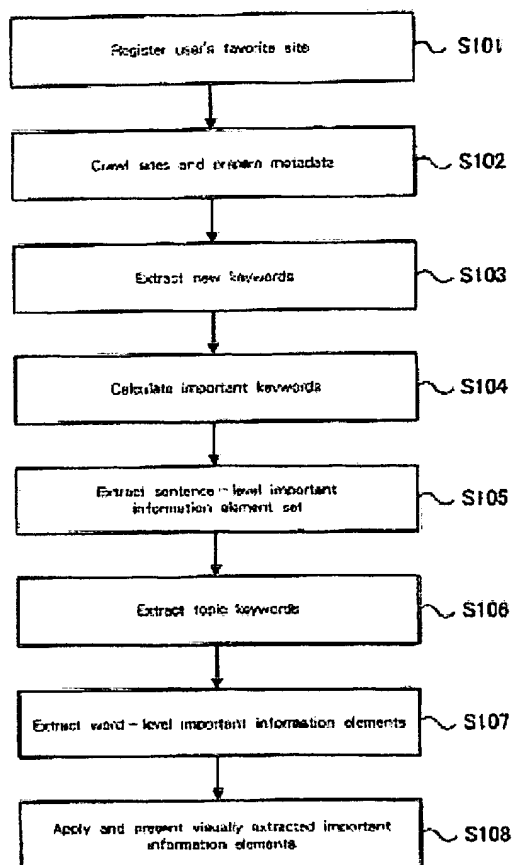
Request for First Extension of Time

Applicants hereby request a first extension of time for reply of two months under 37 CFR 1.137(a)(2) and authorizes the Commissioner to charge our deposit account (09-0441) the requisite fees under 37 CFR 1.17(a)(2).

Remarks/Discussion

Applicant' Invention

Figure 1 provides a high level, visual overview of Applicants' claimed invention:



The high points of the claimed invention are described in the following excerpts of the application:

[0015] To achieve the above objects, according to the present invention, multiple information sources (referred to by, for example, URLs) that are obtained on the Internet or another network and are dynamically changed are periodically observed, and from among the information elements, such as new anchors (links) or text, that are extracted from these information sources, information elements that include the same facts that are referred to on multiple sites or that include a set of important words are extracted and are presented visually. Specifically, according to the present invention, an information rearrangement method for rearranging information obtained from information sources connected via a network comprises: an information collection step of collecting information from a predetermined number of registered sites; an information element extraction step of extracting, from among the collected information, information elements that include the same facts that are referred to at multiple sites; and a display step of displaying the contents of the extracted information elements while changing the display state of the contents in accordance with the number of sites whereat the facts are referred to.

[0016] At the information element extraction step, the information elements that convey the same facts are extracted together with a set of keywords that represent the information elements. This extraction method can be a method for selecting, from the keywords of the information elements included in one set, a keyword having an appearance rate that is equal to or greater than a threshold value. These keywords can be words that are effective for determining the same facts that are included in the information elements, such as new anchors (links) or text, and nouns, such as predetermined proper nouns; these words can be also be verbs.

[0018] According to the present invention, an information rearrangement method comprises: a topic keyword extraction step of extracting a topic keyword that represents the entire set of information elements to be extracted; and a display step of displaying the contents of the extracted information elements, while displaying the extracted topic keyword at a position different from the contents concerning the information elements. For example, only a topic keyword including multiple items is arranged on the upper portion in the display area so that the user can easily see it.

The Art of Record

The primary reference was Published Patent Application 2002/0138487 to Weiss et al. for Method And System For Mapping And Searching The Internet And Displaying The Results In A Visual Form which describes a method for carrying out a search of Web sites according to search criteria. Weiss' method comprises the steps of

- (a) pre-indexing the sites of the Web, including grouping the Web sites according to predefined group-criteria;
- (b) pre-classifying each Web site according to a predefined set of properties;
- (c) pre-visual-formulating each Web site according to its identified properties; and (d) upon searching of Web sites that sustain a search criterion, displaying the formed site results divided into these pre-indexed groups where each site within a group is displayed according to its visual-formulation.

United States Patent 6,638,317 to Nakao for Apparatus And Method For Generating Digest According To Hierarchical Structure Of Topic describes a digest generator for calculating a lexical cohesion degree at each position in a document. The digest generator does this using a plurality of windows having different sizes, and calculates the candidate section of a topic boundary for each topic level corresponding to the size of a window. By unifying the candidate section of different levels, the digest generator detects the topic boundary for each level. Then, based on the relation between a summarization-target topic passage and a long topic passage containing the summarization-target topic passage the digest generator extracts key sentences and generates a digest.

United States Patent 6,289,342 to Lawrence et al. for Autonomous Citation Indexing And Literature Browsing Using Citation Context describes an autonomous citation indexing system. The system can be used as an assistant agent to automate and enhance finding publications in electronic form, including publications located on the world wide web. This is disclosed to be done by parsing citations from papers and identifying citations to the same paper that may differ in syntax. The system also extracts and provides the context of citations to a given paper, allowing a researcher to determine what is published in other papers about a

given paper. Common citations and word or string vector distance similarity are used to find related articles in a search. Figure 3 was cited against claims 8, 14, and 15. Figure 3 is described in the specification as “FIG. 3 shows the results of the query dempster. The articles have been ranked by the number of citations to the articles.”

U.S. Patent 5,875,446 to Brown et al. for System And Method For Hierarchically Grouping And Ranking A Set Of Objects In A Query Context Based On One Or More Relationships describes a system for hierarchically ranking topically relevant objects in an object database. These topically relevant objects are first identified using generally known methods to obtain a set of topically relevant objects (topically relevant set). Parents, (including ancestors), of one or more of the topically relevant objects are identified according to directional structural relationships that the parents have with respect to the topically relevant objects. These objects form a set of structurally relevant objects (structurally relevant set). In some embodiments, the user query identifies one or more of these structural relationships. The topically relevant objects are then organized under one or more of their respective parents to form a hierarchy level of both (topically relevant and structurally relevant) sets of objects. The process can iterate to create more than one hierarchy level.

U.S. Patent 6,665,659 to Logan for Methods And Apparatus For Distributing And Using Metadata Via The Internet, column 1, lines 7-10, “Field of the Invention”¹ is applied against claims 9 and 20, while column 8, lines 21-29² is applied against claim 10. Logan describes a system for selectively distributing information from a multiplicity of Internet resources to a user in a way that makes it easier for the user to quickly identify information of particular interest. The system of Logan et al. employs a server for generating a central library of citations, each containing metadata that describes selected information from a resource

¹ This invention relates to electronic information distribution systems and more particularly to a method for indexing, combining, managing and distributing information via the Internet.

² In addition, the rendering program 141 may include means for accepting a rating value from the user indicating the user's assessment of the quality or usefulness of that citation. If desired, a user may be permitted to compose and attach a user's review of the information specified by a particular citation, or a user-generated cross-reference to a related citation, providing an open-ended forum for review, ranking, comment, and cross-linking to be associated with each citation.

identified by a URL. The server works in conjunction with a client computer which requests information on a topic or topics of interest by supplying preference data to the server. In response, the server delivers a subset of the citations to the client computer which match the preference data from the client. The client computer places this subset of citations in a local store where they may be compared with user requests by matching the metadata in each citation to criteria specified by the user. In addition, the locally stored citations may be sorted into a particular order in response to a user request. The filtered and sorted citations may then be used to present desired information to the user, either by displaying metadata contained in the citation or by using the URL in the citation to fetch relevant information via the Internet from the original resource.

The Office Action of June 12, 2006

The art was applied as shown in the Table below.

The primary reference, Weiss et al,³ was variously applied with Brown⁴, Lawrence⁵, and Logan⁶ as shown in the table below.

³ [0008] Most other search Web applications are highly automated, sending "Spider" programs out on the Web around the clock to collect the text of Web pages. Spiders follow all the links on a page and put all the text into a database. Sometimes a Web site offers both--a search engine and directory capabilities.

[0031] Preferably, the set of properties comprises parameters relating to the site's importance, the nature of the site's owner, the existence of an e-store within the site, the existence of a "chat room" within the site, the existence of a forum within the site, the existence of multimedia file(s) and/or their amount and/or size within the site, the frequent used keywords in the textual data of the site, whether the site is "official", the essence of the site, and/or the amount of information in the site.

[0032] Preferably, the importance of a Web site is a function of the hyperlinks pointing to and from a Web site.

[0106] Marked as 101, is the process that is made by the searching facility that includes Clustering and Labeling. As a result, trees of Clusters are constructed. The Web sites of the Internet are scanned and the titles of the Web sites, the links and the addresses of the pages in which predefined keywords are found are stored in a database. Then, a Clustering algorithm is executed on the collected data in order to determine clusters. Then the detected Clusters are labeled by a Labeling process.

[0137] FIG. 7 schematically illustrates an example of a presentation of the results of the first stage of a search, according to a preferred embodiment of the invention. The search was for the phrase "Charlie's Angels". Optionally, the results are presented in a 2-D map on which the main clusters are displayed as continents: the Sport continent, the Entertainment continent, the Health continent, etc. The Clusters, in which the term "Charlie's Angels" appeared, are marked for the user. Of course alternatively this presentation can be a textual presentation or most preferably 3D presentation.

[0139] After selecting the Entertainment "continent" (the selection being carried out by clicking the selected object), the user is presented with the "countries"—TV series, Movies, Plays, Music, etc. Again, The countries, in which the search subjects have been found, are being marked to the user (see FIG. 8). The size of the "country" is proportional to the number of the Web sites of this entity.

[0147] a Spider program 22, for scanning the Web sites of the Internet;

[0148] a Database 24, for storing the information collected by the Spider program 22;

[0149] an Indexing application 23, for carrying out the clustering, labeling and classification of the Web sites. The indexing is a process, which is carried out independent of the search process, and its purpose is to organize all the Web sites of the Web prior to the search. For example, the indexing concerns organizing all the Web sites in clusters, classifying the Web sites according to predetermined properties, etc.; and

[0150] a Seeker application program/server 28 for interacting with the users 25, carrying out the search (by the appropriate queries to database 24) and for sending the results to the users 25 (usually as Web pages, which usually perform a visual presentation of user's Web browser).

[0155] According to an embodiment of the present invention, the Indexing 23 comprises the activities of Clustering, Labeling and Classification of the Web sites according to the predefined attributes, as described above. A Spider program 22 scans the Web sites of the Internet. The found Web sites are added to a database 24. By implementing Clustering method(s) a tree of Clusters is obtained. The gathered information (tree of Clusters, and the list of Web sites and their classification) is stored in database 24.

[0158] Then a query is posted from the system to database 24, and the results of the query are presented to the user. This stage is carried out by a Seeker program 32. The results of the search may be presented in a textual form or, but preferably in a graphical form described above (marked as 33). If the user is not satisfied with the search results, then the system may interview the user in order to focus the search, and the system posts a new query to the database 24.

⁴ Column 4, lines 43-45:

An object of this invention is a system and method that generates a hierarchical grouping of topically and structurally relevant objects in a query context.

Column 16, lines 13-29:

The result of step 900 is a display of ranked hierarchies where children are shown grouped and indented under their parent. An example of such a display is shown in FIG. 13. FIG. 13 shows a sample output result of the system. The Figure shows the result of iterating step 635 once, such that a two level hierarchy is generated. The original topically relevant objects supplied in step 610 are displayed indented as 1320. The structurally relevant parent objects found after one iteration of step 635 are displayed non-indented as 1310. The parent objects 1310 form the next level of the hierarchical view, provide navigational starting points for browsing the relevant objects, and group the topically relevant child objects 1320. The display provides the end user with insight into the structure of the object collection being searched. Attributes for each of the objects shown in the display are obtained from the Object Catalog 210 and Attribute Tables 250.

⁵ For Figure 3, and its teaching of an "autonomous citation index system."

⁶ Column 1, lines 7-10:

This invention relates to electronic information distribution systems and more particularly to a method for indexing, combining, managing and distributing information via the Internet.

Claim #	Objections and §112	§103 – Weiss + Nakao	§103 – Weiss + Nakao, Brown	§103 – Weiss + Nakao, Lawrence	§103 – Weiss + Nakao, Logan
1	X	X			
2	X	X			
4	X		X		
5	X		X		
6	X	X			
7	X	X			
8	X			X	
9	X				X
10	X				X
13		X			
14				X	
15				X	
16			X		
17			X		
18		X			
19		X			
20					X
21					X
24	X	X			
25	X		X		
26	X	X			

Column 5, lines 3-10:

The metadata which describes a particular resource need not be derived directly from that resource; for example, external resources which link to or which review a given resource may be used as a source of metadata which describes the given resource. By way of example, reviews or comments on a given Web site may be analyzed to develop an attribute value which quantifies the degree to which such comments and reviews were favorable or unfavorable.

Remarks/Arguments

With reference to the Office Action mailed June 12, 2006, Applicants offer the following remarks and argument.

Status of the Claims

Applicants have previously canceled claims 11-13 and 27-28 as drawn to a non-elected invention. The pending claims are claims 1, 2, 4-10 and 13-21, and 24-26, being the claims of Group 1.

The remaining claims, or the base claims on which they depend, have been substantially amended to positively recite and characterize the “extraction” process. In that regard, claims 1⁷ and 16⁸ are exemplary.

⁷ (Currently Amended) 1. An information rearrangement method for rearranging information obtained from information sources connected via a network comprising: an information collection step of collecting information from a predetermined number of registered sites; an information element extraction step of extracting, from among said collected information, information elements that include the same facts that are referred to at multiple sites, said information element extraction step comprising selecting, from keywords of information elements included in one set, a keyword having an appearance rate that is equal to or greater than a threshold value, said keywords comprising words that are keywords searchable, rankable, extractable, representations of information elements for determining the same facts included in the information elements, and chosen from the group consisting of anchors, links, text, nouns, predetermined proper nouns, and predetermined verbs; and a display step of displaying the contents of said extracted information elements while changing the display state of said contents in accordance with the number of sites whereat said facts are referred to; said step of displaying comprising extracting a set of important information elements on a sentence level from a group composed of a predetermined number of sites, and folding the display for the same sets of important information elements.

⁸ (Previously Amended) 16. An information processing apparatus for rearranging information obtained from information sources connected via a network comprising: information collection means for collecting information from a predetermined number of registered sites; information element extraction means for extracting, from among said collected information, information elements that include the same facts that are referred to at multiple sites, the information element extraction means comprising a keyword extraction means for extracting keywords based on keywords for metadata stored in a metadata; a keyword importance level calculation means for calculating a keyword importance; a sentence-level information element extraction means for calculating a set of sentence-level information elements, a topic keyword extraction means for extracting from the entire set of information elements extracted by the sentence-level information elements extraction means, and a word-level information element extraction means, for extracting information elements in which a combination of keywords that are obtained by the keyword importance level calculation means appear, topic keyword extraction means for extracting a topic keyword that

Discussion

Objections and Rejections under 35 USC 112

Applicants have amended the claims to obviate the objections and the rejections under 35 USC 112.

Specifically, Applicants have canceled “effective” and replaced with the term “keyword.”

Rejections Under 35 USC 103

Applicants’ claims define an invention that is allowable over. Claim 1 is exemplary, and the discussion herein will be with respect to Claim 1.

Claim 1 was amended is as follows:

(Currently Amended) 1. An information rearrangement method for rearranging information obtained from information sources connected via a network comprising: an information collection step of collecting information from a predetermined number of registered sites; an information element extraction step of extracting, from among said collected information, information elements that include the same facts that are referred to at multiple sites, said information element extraction step comprising selecting, from keywords of information elements included in one set, a keyword having an appearance rate that is equal to or greater than a threshold value, said keywords comprising words that are keywords searchable, rankable, extractable, representations of information elements for determining the same facts included in the information elements, and chosen from the group consisting of anchors, links, text, nouns, predetermined proper nouns; and predetermined verbs; and a display step of displaying the contents of said extracted information elements while changing the display state of said contents in accordance with the number of sites whereat said facts are referred to; said step of displaying comprising extracting a set of important information elements on a sentence level from a group composed of a predetermined number of sites, and folding the display for the same sets of important information elements.

represents the entire set of information elements to be extracted, said topic keyword extraction means comprising a representative keyword extraction means, a set representative keyword extraction means, and a topic keyword collection means, whereby results extracted by sentence-level important information element extraction means, topic keyword extraction means, and word-level important information element extraction means are stored in an important information element; and display means for displaying the contents of said extracted information elements, while displaying said extracted topic keyword at a position different from the contents concerning said information elements.

While Applicants and Weiss are both addressed to searching the internet, there are significant differences. This can be seen at a high level by comparing Applicants' claim preamble, "An information rearrangement method for rearranging information obtained from information sources connected via a network comprising:" with Weiss's Summary, i.e.,

[0029] In one aspect, the present invention is directed to a method for carrying out a search of Web sites according to a search criteria, comprising: pre-indexing the sites of the Web, including grouping the Web sites according to predefined group-criteria; pre-classifying each Web site according to a predefined set of properties; pre-visual-formulating each Web site according to its identified properties; and upon searching of Web sites that sustain a search criterion, displaying the formed site results divided into the pre-indexed groups wherein each site within a group is displayed according to its visual-formulation.

That is, "a search of web sites" where the search is "according to a search criteria" with "pre-indexing the sites of the Web, including grouping the Web sites according to predefined group-criteria; pre-classifying each Web site according to a predefined set of properties; pre-visual-formulating each Web site according to its identified properties; and upon searching of Web sites that sustain a search criterion, displaying the formed site results divided into the pre-indexed groups wherein each site within a group is displayed according to its visual-formulation."

Going into a comparison of Applicants' claim language to Weiss' disclosure, Applicants' claim 1 calls for "selecting, from keywords of information elements included in one set, a keyword having an appearance rate that is equal to or greater than a threshold value, said keywords comprising words that are keywords for determining the same facts included in the information elements, and chosen from the group consisting of anchors, links, text, nouns, predetermined proper nouns; and predetermined verbs;" in contradistinction to Weiss's disclosure of "[0031] Preferably, the set of properties comprises parameters relating to the site's importance, the nature of the site's owner, the existence of an e-store within the site, the existence of a "chat room" within the site, the existence of a forum within the site, the existence of multimedia file(s) and/or their amount and/or size within the site, the frequent used keywords in the textual data of the site, whether the site is "official", the essence of the site, and/or the amount of information in the site."

By way of comparison and contrast, Weiss discloses a less fine (more coarse) measure, i.e.,

“[0032] Preferably, the importance of a Web site is a function of the hyperlinks pointing to and from a Web site.

“[0033] Preferably, the amount of information in a Web site is determined according to the number of characters, and/or the number of words, and/or the number of bytes included within the Web site.”

This disclosure of the number of hyperlinks into or out of site and the size of the site is different then the number of identified keywords in the site. Size and number of hyperlinks is clearly different from Applicants’ use of “keywords” as described above⁹.

And while Weiss’s mentions “keywords” the context is clearly different¹⁰ and Weiss’s description of “clustering” is “on a sentence level”¹¹ while other cited parts of Weiss et al. describe “Indexing”, i.e., [0146]-[0150] and output [0158]. Moreover, Applicants have amended their claims to characterize the keywords as *“keywords comprising words that are searchable, rankable, extractable, representations of information elements for determining the same facts included in the information elements, and chosen from the group consisting of anchors, links, text, nouns, predetermined proper nouns; and predetermined verbs.”*

⁹ “selecting, from keywords of information elements included in one set, a keyword having an appearance rate that is equal to or greater than a threshold value, said keywords comprising words that are keywords for determining the same facts included in the information elements, and chosen from the group consisting of anchors, links, text, nouns, predetermined proper nouns; and predetermined verbs;”

¹⁰ [0045] In another aspect, the invention is directed to a method for finding sub-groups having a common basis in a set of Web sites, comprising: clustering the set of Web sites by determining groups having a common basis by their being related by hyperlink(s) pointing to and from each of the Web sites; and labeling the determined groups by analyzing their content.

[0046] Preferably, the analyzing is carried out by detecting keywords frequently used in a determined group. (emphasis added)

¹¹ [0137] FIG. 7 schematically illustrates an example of a presentation of the results of the first stage of a search, according to a preferred embodiment of the invention. The search was for the phrase “Charlie’s Angels”. Optionally, the results are presented in a 2-D map on which the main clusters are displayed as continents: the Sport continent, the Entertainment continent, the Health continent, etc. The Clusters, in which the term “Charlie’s Angels” appeared, are marked for the user. Of course alternatively this presentation can be a textual presentation or most preferably 3D presentation.

Nakao does not overcome the deficiencies of Weiss. Specifically, Claim 10 of Nakao¹² recites such limitations as

keyword extracting,

calculating a likelihood ratio based on a comparison of a use frequency of the word in the process target topic passage with a use frequency of the word in a longer topic passage including the process target topic passage and comparing the likelihood ratio to a predetermined threshold value,

extracting a keyword from the process target topic passage when the likelihood ratio is greater than the predetermined threshold value;

generating a digest according to a use condition of said keyword; and outputting said digest.

But, what Nakao fails to show are the claim limitations of

(the) keywords comprising words that are keywords for determining the same facts included in the information elements,

chosen from the group consisting of anchors, links, text, nouns, predetermined proper nouns; and predetermined verbs;

a display step of displaying the contents of said extracted information elements while changing the display state of said contents in accordance with the number of sites whereat said facts are referred to;

¹² 10. A digest generator apparatus, comprising:

a keyword extracting unit evaluating whether or not a word is characteristic of a process target topic passage in a document by calculating a likelihood ratio based on a comparison of a use frequency of the word in the process target topic passage with a use frequency of the word in a longer topic passage including the process target topic passage and comparing the likelihood ratio to a predetermined threshold value, and extracting a keyword from the process target topic passage when the likelihood ratio is greater than the predetermined threshold value;

a generating unit generating a digest according to a use condition of said keyword; and an outputting unit outputting said digest.

and said step of displaying comprising extracting a set of important information elements on a sentence level from a group composed of a predetermined number of sites, and folding the display for the same sets of important information elements.

These material limitations are neither taught nor suggested by Nakao.

Turning first to claims 1 and 24, the claim limitation *“rearrangement method for rearranging information obtained from information sources connected via a network comprising: an information collection step of collecting information from a predetermined number of registered sites;”* was rejected on paragraphs [0147]¹³ and [0148]¹⁴ of Weiss. The encyclopedic web site www.whatis.com defines a “spider as

“A spider is a program that visits Web sites and reads their pages and other information in order to create entries for a search engine index. The major search engines on the Web all have such a program, which is also known as a “crawler” or a “bot.” Spiders are typically programmed to visit sites that have been submitted by their owners as new or updated. Entire sites or specific pages can be selectively visited and indexed. Spiders are called spiders because they usually visit many sites in parallel at the same time, their “legs” spanning a large area of the “web.” Spiders can crawl through a site’s pages in several ways. One way is to follow all the hypertext links in each page until all the pages have been read.”

Next, the claim limitation of *“an information element extraction step of extracting, from among said collected information, information elements that include the same facts that are referred to at multiple sites,”* was said to be taught by Weiss numbered paragraphs [0149]¹⁵ and [0150]¹⁶. Numbered paragraphs [0149] and [0150] describe the steps of “indexing” (organizing the web sites in clusters and classifying the web sites according to predetermined properties) and searching the “indexed” web sites. As described in numbered paragraph [0149], “indexing” has the purpose of “organiz[

¹³ [0147] a Spider program 22, for scanning the Web sites of the Internet;

¹⁴ [0148] a Database 24, for storing the information collected by the Spider program 22;

¹⁵ [0149] an Indexing application 23, for carrying out the clustering, labeling and classification of the Web sites. The indexing is a process, which is carried out independent of the search process, and its purpose is to organize all the Web sites of the Web prior to the search. For example, the indexing concerns organizing all the Web sites in clusters, classifying the Web sites according to predetermined properties, etc.;

¹⁶ [0150] a Seeker application program/server 28 for interacting with the users 25, carrying out the search (by the appropriate queries to database 24) and for sending the results to the users 25 (usually as Web pages, which usually perform a visual presentation of user’s Web browser).

ling all the Web sites of the Web prior to the search. For example, the indexing concerns organizing all the Web sites in clusters, classifying the Web sites according to predetermined properties, etc”

This is in contrast to Applicants’ description of “extracting,” as described in numbered paragraph [0055], i.e., *”This metadata is obtained by extracting, from the contents referred to by a URL, elements that convey information, and consists of a link and a text portion, or continuous text portions.”* and numbered paragraphs [0056] and [0057], i.e.,

[0056] At each registered site, keywords that have appeared in a version during a designated period and a preceding period are counted, and new words (keywords) are extracted by weighting a set of keywords (step 103). Thereafter, important keywords are calculated by using the sum of the ratios of new words to the keywords (step 104). Then, the importance level obtained from a corresponding keyword is defined as the importance level of an information element, and the information elements are sorted in the order of their importance levels. Thereafter, a set of information elements (a set of sentence-level important information elements) that convey the same facts are extracted from the information element sets at the individual sites (step 105), and the importance level is provided in accordance with the number of sites that are referred to. Then, in order to extract the information elements that convey the same facts, a keyword having a category that is extracted from the text portion of the information element is used, and whether a proper noun is included is important for determining whether the information elements include the same facts.

[0057] A set of keywords (representative keywords) are extracted from the individual sentence-level important information elements that are extracted. Then, a keyword (topic keyword) that is representative of the entire set of information elements is extracted from the representative keywords (step 106). The importance levels can be determined by using the number of important information element sets that appear and the ratio of the arrival of new keywords. Further, a determination reference as to whether many words are referred to at multiple sites is employed to extract the information elements that are assumed to be important (word-level important information elements) (step 107). The important information elements extracted at the individual steps are applied to a group of sites or a single site, and are presented visually (step 108). When the information elements are applied for a single site, they are identified and presented visually by changing a group of sites into an arbitrary group. By using the above described important information element extraction method, a user can extract important information from a mass of data and read it.

The “extracting” contemplated, described, and claimed by Applicants is materially different from the “indexing” of Weiss et al., and Weiss et al. neither teaches nor suggests this “extracting.”

Only in discussing the claimed element of *“a display step of displaying the contents of said extracted information elements while changing the display state of said*

contents in accordance with the number of sites whereat said facts are referred to” is Weiss’ numbered paragraph [0158]¹⁷

Next, the Office Action recites that the claim limitation *“comprising extracting a set of important information elements on a sentence level from a group composed of a predetermined number of sites, and folding the display for the same sets of important information elements.”* is said to be taught by Weiss numbered paragraphs [0137]¹⁸, [0139]¹⁹ as is the use of keywords, numbered paragraphs [0031]²⁰, [0046]²¹.

The teachings relied upon to show “keywords” fail to show the claim element of *“keywords comprising words that are searchable, rankable, extractable, representations of information elements for determining the same facts included in the information elements, and chosen from the group consisting of anchors, links, text, nouns, predetermined proper nouns; and predetermined verbs”*

¹⁷ [0158] Then a query is posted from the system to database 24, and the results of the query are presented to the user. This stage is carried out by a Seeker program 32. The results of the search may be presented in a textual form or, but preferably in a graphical form described above (marked as 33). If the user is not satisfied with the search results, then the system may interview the user in order to focus the search, and the system posts a new query to the database 24.

¹⁸ [0137] FIG. 7 schematically illustrates an example of a presentation of the results of the first stage of a search, according to a preferred embodiment of the invention. The search was for the phrase “Charlie’s Angels”. Optionally, the results are presented in a 2-D map on which the main clusters are displayed as continents: the Sport continent, the Entertainment continent, the Health continent, etc. The Clusters, in which the term “Charlie’s Angels” appeared, are marked for the user. Of course alternatively this presentation can be a textual presentation or most preferably 3D presentation.

¹⁹ [0139] After selecting the Entertainment “continent” (the selection being carried out by clicking the selected object), the user is presented with the “countries”—TV series, Movies, Plays, Music, etc. Again, The countries, in which the search subjects have been found, are being marked to the user (see FIG. 8). The size of the “country” is proportional to the number of the Web sites of this entity.

²⁰ [0031] Preferably, the set of properties comprises parameters relating to the site’s importance, the nature of the site’s owner, the existence of an e-store within the site, the existence of a “chat room” within the site, the existence of a forum within the site, the existence of multimedia file(s) and/or their amount and/or size within the site, the frequent used keywords in the textual data of the site, whether the site in “official”, the essence of the site, and/or the amount of information in the site.

²¹ [0046] Preferably, the analyzing is carried out by detecting keywords frequently used in a determined group.

Weiss' recitation of numbered paragraphs [0031] and [0046], with recitations of "the frequent used keywords in the textual data of the site, whether the site in "official", the essence of the site, and/or the amount of information in the site." and "Preferably, the analyzing is carried out by detecting keywords frequently used in a determined group" does not describe Applicants claimed method of extracting information and using keywords, i.e.,

- collecting information from a predetermined number of registered sites;*
- an information element extraction step of extracting, from among said collected information*
- said information element extraction step comprising selecting, from keywords of information elements included in one set, a keyword having an appearance rate that is equal to or greater than a threshold value,*
- said keywords comprising words that are searchable, rankable, extractable, representations of information elements*
- for determining the same facts included in the information elements, and*
- chosen from the group consisting of anchors, links, text, nouns, predetermined proper nouns; and predetermined verbs; and*
- a display step of displaying the contents of said extracted information elements while changing the display state of said contents in accordance with the number of sites whereat said facts are referred to;*
- said step of displaying comprising extracting a set of important information elements on a sentence level from a group composed of a predetermined number of sites, and*
- folding the display for the same sets of important information elements.*

As conceded in the Office Action, Weiss does not teach checking if the appearance rate of the keyword is equal to or greater than a threshold value. The citation of Nakao, column 41, lines 5-20 (Nakao, claim 10)²² introduces the further complexity²³ of a "likelihood value" which is neither claimed nor disclosed by Applicants.

²² 10. A digest generator apparatus, comprising: a keyword extracting unit evaluating whether or not a word is characteristic of a process target topic passage in a document by calculating a likelihood ratio based on a comparison of a use frequency of the word in the process target topic passage with a use frequency of the word in a longer topic passage including the process target topic passage and comparing the likelihood ratio to a predetermined threshold value, and extracting a keyword from the process target topic passage when the likelihood ratio is greater than the predetermined threshold

Moreover neither Weiss nor Nakao, either alone or in combination, teach or suggest a sequence of:

- an information element extraction step of extracting, from among said collected information*
- said information element extraction step comprising selecting, from keywords of information elements included in one set, a keyword having an appearance rate that is equal to or greater than a threshold value,*
- said keywords comprising words that are searchable, rankable, extractable, representations of information elements*
- for determining the same facts included in the information elements, and*

value; a generating unit generating a digest according to a use condition of said keyword; and an outputting unit outputting said digest.

23 Here, if the use frequency (number of uses) of a certain content word t in a topic section b exceeds an expectation, and the logarithmic likelihood ratio L of the following equation exceeds a given threshold (χ^2 value corresponding to a statistically significant standard), the content word t is judged to be characteristic of the topic section b .

##EQU4##

In equation (19), $F_{sub.bt}$ indicates the use frequency of a word t in a topic section b , $F_{sub.at}$ indicates the use frequency of the word t in the parent topic section a of the topic section b , and $E(F_{sub.bt})$ indicates the expectation of the use frequency of the word t in the topic section b . $E(F_{sub.bt})$ can be calculated by multiplying the use density (use probability) of the word t in the parent topic section a by the size of the topic section b . Here, the use density of a word in a certain section means a ratio between the use frequency of the word and the size of the section.

The L of the equation (19) is the value of a likelihood ratio test on whether or not the use probability of a word t is independent of a distinction between a topic section b and areas other than the topic section, which indicates that the greater this value is, the more the use probability of the word is dependent on the distinction. Since the degree of freedom v of L is 1, when the significant standard is 10%, 5% and 1%, it is sufficient that the threshold is set to 6.63490, 7.87994 and 10.8276, respectively. Instead of using a threshold, several higher-order words can also be extracted as keywords in the descending order of L .

However, if a topic section based on a maximum window width $w_{sub.0}$ and a topic section based on the basic window width $w_{sub.1}$ are identical, or if a topic section of the basic window width $w_{sub.1}$ occupies most of a topic section based the window width $w_{sub.0}$, such a test method cannot always work well. For this reason, when the size of the direct higher-order topic section of b (i.e., a topic section including b of a window width $w_{sub.0}$) is less than twice the size of b , an entire document is used for the parent topic section.

- chosen from the group consisting of anchors, links, text, nouns, predetermined proper nouns; and predetermined verbs; and*
- a display step of displaying the contents of said extracted information elements while changing the display state of said contents in accordance with the number of sites whereat said facts are referred to;*
- said step of displaying comprising extracting a set of important information elements on a sentence level from a group composed of a predetermined number of sites, and*
- folding the display for the same sets of important information elements.*

and more specifically, no argument can be made that either Weiss or Nakao teaches or suggests the claimed limitations of:

- the keywords comprising words that are searchable, rankable, extractable, representations of information elements*
- the keywords determining the same facts included in the information elements, and*
- the keywords are chosen from the group consisting of anchors, links, text, nouns, predetermined proper nouns; and predetermined verbs; and*
- the step of displaying comprises extracting a set of important information elements on a sentence level from a group composed of a predetermined number of sites.*

Thus claims 1 and 24 and all of the claims ultimately dependent thereon are clearly allowable to Applicants.

Claim 2, dependent on claim 1, is properly allowable notwithstanding the citation of Weiss numbered paragraph [0137] because of its dependency from allowable claim 1.

Independent Claims 4, 6, 9, 13, 16, 18, 20, 24, and 26 have all been amended to contain the above limitations added to claim 1, and the discussion with respect to claim 1 applies to claims 4, 6, 9, 13, 16, 18, 20, 24, and 26, and all of the claims dependent thereon.

Claims 4, 13, 18, 24, and 26 contain substantially the same limitations as claim 1, and the same art was cited against them. Thus, the same arguments apply

With respect to claim 9²⁴, the passage in Logan, column 5, lines 3-10²⁵, Logan neither teaches or suggests

“determining whether, of said information elements extracted from said multiple sites, there are relevant information elements that convey the same facts as sentence-level information elements that constitute an arbitrary web page; and when said relevant information elements that include the same facts as said sentence-level information elements are present in said information elements obtained from said multiple sites, adding remark information to said sentence-level information elements to provide information concerning said arbitrary web page.”

but, instead states that “the metadata which describes a particular resource need not be derived directly from that resource; for example, external resources which link to or which review a given resource may be used as a source of metadata which describes the given resource.” This is a use of externalities, for example Logan describes “reviews or comments on a given Web site may be analyzed to develop an attribute value which quantifies the degree to which such comments and reviews were favorable or unfavorable.”

With respect to independent claims 16²⁶ and 25²⁷ the cited clause at column 4, lines 43-45 of Brown²⁸ neither teaches nor suggests the claimed combination. Instead it teaches that an

²⁴ (Previously Amended) 9. An information rearrangement method comprising the steps of: extracting information elements from multiple sites; said step of extracting information elements comprising selecting, from keywords of information elements included in one set, a keyword having an appearance rate that is equal to or greater than a threshold value, said keywords comprising words that are ~~keywords-effective~~ for determining the same facts included in the information elements, and chosen from the group consisting of anchors, links, text, nouns, predetermined proper nouns; and predetermined verbs, and determining whether, of said information elements extracted from said multiple sites, there are relevant information elements that convey the same facts as sentence-level information elements that constitute an arbitrary web page; and when said relevant information elements that include the same facts as said sentence-level information elements are present in said information elements obtained from said multiple sites, adding remark information to said sentence-level information elements to provide information concerning said arbitrary web page.

²⁵ The metadata which describes a particular resource need not be derived directly from that resource; for example, external resources which link to or which review a given resource may be used as a source of metadata which describes the given resource. By way of example, reviews or comments on a given Web site may be analyzed to develop an attribute value which quantifies the degree to which such comments and reviews were favorable or unfavorable.

²⁶ (Currently Amended) An information processing apparatus for rearranging information obtained from information sources connected via a network comprising: information collection means for collecting information from a predetermined number of registered sites; information element extraction means for

object of the Brown invention is a system and method that generates a hierarchical grouping of topically and structurally relevant objects in a query context. By way of contrast the claims recite an *information processing apparatus for rearranging information* obtained from information sources connected via a network. The claimed apparatus comprises: information collection means for collecting information from a predetermined number of registered sites; information element extraction means for extracting, from among said collected information, information elements that include the same facts that are referred to at multiple sites, the *information element extraction means comprising a keyword extraction means* for extracting keywords based on keywords for metadata stored in a metadata database; a keyword importance level calculation means for calculating a keyword importance; a sentence-level information element extraction means for calculating a set of sentence-level information elements, a topic keyword extraction means for extracting from the entire set of information

extracting, from among said collected information, information elements that include the same facts that are referred to at multiple sites, the *information element extraction means comprising a keyword extraction means* for extracting keywords based on keywords for metadata stored in a metadata database; a keyword importance level calculation means for calculating a keyword importance; a *sentence-level information element extraction means* for calculating a set of sentence-level information elements, a topic keyword extraction means for extracting from the entire set of information elements extracted by the sentence-level information elements extraction means, and a word-level information element extraction means, for extracting information elements in which a combination of keywords that are obtained by the keyword importance level calculation means appear, topic keyword extraction means for extracting a topic keyword that represents the entire set of information elements to be extracted, said topic keyword extraction means comprising a representative keyword extraction means, a set representative keyword extraction means, and a topic keyword collection means, whereby results extracted by sentence-level important information element extraction means, topic keyword extraction means, and word-level important information element extraction means are stored in an important information element; and display means for displaying the contents of said extracted information elements, while displaying said extracted topic keyword at a position different from the contents concerning said information elements.

²⁷ (Currently Amended) 25. A storage medium on which a computer-readable program is stored, which permits a computer to perform: a process for collecting information from a predetermined number of registered sites; a process for extracting, from among said collected information, information elements that include the same facts that are referred to at multiple sites, said information element extraction processing step comprising selecting, from keywords of information elements included in one set, a keyword having an appearance rate that is equal to or greater than a threshold value, said keywords comprising words that are keywords effective for determining the same facts included in the information elements, and chosen from the group consisting of anchors, links, text, nouns, predetermined proper nouns; and predetermined verbs;[[[;]]] a process for extracting a topic keyword that represents the entire set of information elements to be extracted; and a process for displaying the contents of said extracted information elements, while displaying said extracted topic keyword at a position different from the contents concerning said information elements.

²⁸ An object of this invention is a system and method that generates a hierarchical grouping of topically and structurally relevant objects in a query context.

elements extracted by the sentence-level information elements extraction means, and a word-level information element extraction means, for extracting information elements in which a combination of keywords that are obtained by the keyword importance level calculation means appear, topic keyword extraction means for extracting a topic keyword that represents the entire set of information elements to be extracted, said topic keyword extraction means comprising a representative keyword extraction means, a set representative keyword extraction means, and a topic keyword collection means, whereby results extracted by sentence-level important information element extraction means, topic keyword extraction means, and word-level important information element extraction means are stored in an important information element; and display means for displaying the contents of said extracted information elements, while displaying said extracted topic keyword at a position different from the contents concerning said information elements.

Thus, the claimed combination is not taught or suggested by Nakao alone, Logan alone, Brown alone, Lawrence alone, Weiss alone, or any combination of Weiss with Nakao, Lawrence, Brown, or Logan.

Conclusion

It is respectfully submitted that the pending claims address the points raised in the Office Action of January 30, 2006 and describe an invention that is properly allowable to the Applicants.

If any issues remain unresolved despite the present amendment, the Examiner is requested to telephone Applicants' Attorney at the telephone number shown below to arrange for a telephonic interview before issuing another Restriction Requirement or Advisory Action.

Applicants would like to take this opportunity to thank the Examiner for a thorough, complete, exhaustive, and competent examination and for courtesies extended to Applicants' Attorney.

Respectfully Submitted


Certificate of Mailing

I hereby certify that this paper (along with any paper referred to as being attached or enclosed) is being deposited with the United States Postal Service on the date shown below with sufficient postage as Certified Priority Mail (Certified Label 7006 0810 0004 8352 6402) in an envelope addressed to the Commissioner for Patents, Mail Stop NFA, PO Box 1450, Alexandria Virginia, 22313-1450

Date of deposit: November 13, 2006

Person mailing paper: Richard M. Goldman

Signature: 


Richard M. Goldman, Reg. # 25,585
371 Elan Village Lane, Suite 208
San Jose, CA 95134
Voice: 408-324-0716
Fax: 408-324-0672
E-mail: goldmanptn@aol.com